

CLAIMS

- 1 1. A method of operating a cache coherency mechanism for a distributed computer sys-
2 tem that includes multiple processors, the method including the steps of:
 - 3 A. determining which processors have copies of data of interest;
 - 4 B. determining paths through various system switching devices on routes from an
5 associated home node to the processors that have copies of the data of interest;
 - 6 C. encoding information that is indicative of the paths into one or more masks;
 - 7 D. when the data of interest is the subject of an update operation, producing at
8 the home node an invalidate message that includes the masks;
 - 9 E. at the switching devices, decoding the applicable masks and routing the in-
10 validate message over the paths indicated by the decoded information; and
 - 11 F. at switching devices that connect to the processors, forwarding the invalidate
12 message to the processors that have copies of the data of interest.
- 1 2. The method of claim 1, wherein the step of encoding includes setting bits in a given
2 mask to indicate paths from a corresponding switching device to a next switching de-
3 vice on the routes to the processors.
- 1 3. The method of claim 2, wherein the step of encoding further includes setting the bits
2 to correspond to a combination of the paths through a plurality of switching devices.
- 1 4. The method of claim 2, wherein the step of encoding further includes setting bits that
2 correspond to ports of the respective switching devices.
- 1 5. The method of claim 1 further including in the step of encoding
 - 2 a. separately encoding into a first mask information relating to paths through the
3 switching device that is associated with the home node,
 - 4 b. encoding into one or more second masks information relating to paths through
5 the switching devices that connect to the switching device of step a,

- 6 c. encoding into one or more additional masks information relating to paths
7 through the switching devices that connect to the switching devices of the
8 previous step, and
9 d. repeating step c for paths through additional switching devices.
- 1 6. A method of operating a cache coherency mechanism for a distributed computer sys-
2 tem that includes multiple processors which are interconnected by layers of switching
3 devices, the method including the steps of:
- 4 A. determining which processors have copies of data of interest;
5 B. determining paths through various system switching devices on routes from a
6 home node to the processors that have copies of the data of interest;
7 C. encoding information that is indicative of the paths through a highest layer of
8 the system into a first mask;
9 D. encoding information that is indicative of the paths through a next highest
10 layer of the system into a next mask;
11 E. repeating step D for the remaining layers of the system;
12 F. when the data of interest is the subject of an update operation, producing at
13 the home node an invalidate message that includes the masks;
14 G. at the switching devices in the highest layer, decoding the first mask and
15 routing the invalidate message over the indicated paths;
16 H. at the switching devices in the remaining layers, decoding the corresponding
17 masks and routing the invalidate message over the indicated paths through the
18 layers; and
19 I. at switching devices that connect to the processors of interest, forwarding the
20 invalidate message to the processors.
- 1 7. The method of claim 6, wherein the steps of encoding include setting bits in a given
2 mask to indicate one or more paths from a corresponding switching device to one or
3 more switching devices in a next layer of the system.

- 1 8. The method of claim 6, wherein the steps of encoding further include setting the bits
2 to correspond to a combination of the paths through a plurality of switching devices
3 in a given layer of the system.
- 1 9. The method of claim 7, wherein the steps of encoding further include setting bits that
2 correspond to ports of the switching devices.
- 1 10. The method of claim 6 wherein the highest layer includes one or more switching de-
2 vices that receive messages from the home node.
- 1 11. A distributed computer system including:
2 A. a plurality of processors, with one or more processors designated as home
3 nodes;
4 B. a plurality of switching devices that interconnect the processors;
5 C. one or more encoders for encoding into one or more masks information relat-
6 ing to paths through the switching devices from an associated home node to
7 the processors that have data of interest;
8 D. a cache coherency directory with entries for data of interest, the directory in-
9 cluding in a given entry
10 a. information that identifies the owner of the data, and
11 b. one or more associated masks; and
12 E. one or more decoders at the switching devices, the decoder in a given switching
13 device decoding an associated mask to set paths through the switching device for
14 messages directed from the home node to processors that have copies of the asso-
15 ciated data of interest.
- 1 12. The distributed computer system of claim 11 wherein
2 i. the plurality of switching devices are organized into layers with one or more
3 switching devices in a highest layer connected to transmit messages from the
4 home node, one or more switching devices in a next highest layer connected to
5 transmit messages from the one or more switching devices in the highest layer to

6 the switching devices in a lower layer, one or more switching devices in lower
7 layers connected to transmit messages from the switching devices in preceding
8 levels to switching devices in subsequent layers, and one or more switching de-
9 vices in a lowest level connected to transmit messages to the processors, and
10 ii. the masks relate, respectively, to paths through the switching devices in the asso-
11 ciated layers.

1 13. The distributed computer system of claim 12 wherein one or more of the masks relate
2 to combinations of the paths through the switching devices in the associated layers.

1 14. The distributed computer system of claim 11 wherein a given home node produces
2 messages directed to processors that have copies of data of interest and includes in the
3 messages the associated masks.

1 15. The distributed computer system of claim 12 wherein the switching devices are
2 switches and the masks designate port of the associated switches.

1 16. The distributed computer system of claim 14 wherein the switches that connect to the
2 processors use local routing information to provide the messages to the associated
3 processors that have copies of the data of interest.

1 17. The distributed computer system of claim 14 wherein the switches that connect to the
2 processors locally broadcast the messages to the associated processors.